#7

Kongeriget Danmark

Patent application No.: PA 2000 00705

Date of filing: 28 April 2000

Applicant: Christian E. Elling
Viktoriagade 3,1.th.
DK-1655 København V

Thue W. Schwartz
Steen Blichers Vej 13, st.
DK-2000 Frederiksberg

The attached photocopy is a true copy of the following document:

- The specification, claims and drawing as filed with the application on the filing date indicated above.

Patent- og Varemærkestyrelsen ° Kongeriget Danmark °

Patent- og
Varemærkestyrelsen
Erhvervsministeriet

TAASTRUP 17 January 2001

Lizzi Vester
Head of Section

A METHOD OF IDENTIFYING LIGANDS OF PROTEINS

FIELD OF INVENTION

The present invention relates to a method of identifying compounds capable of interacting with proteins in such a way as to modify the biological activity thereof. The invention further relates to methods of identifying compounds acting as ligands of membrane proteins involving the introduction of metal ion binding sites into membrane proteins, including a method of identifying compounds that bind to orphan receptors.

BACKGROUND OF THE INVENTION

Membrane proteins constitute a numerous and varied group of proteins whose function is, *inter alia*, to mediate intercellular communication and communication between the cell exterior and the interior by transducing chemical signals across cell membranes. Membrane proteins are for instance receptors and ion channels to which specific chemical messengers termed ligands are bound resulting in the generation of a signal giving rise to a specific intracellular response (this process is known as signal transduction). Most membrane proteins are anchored in the cell membrane by a sequence of amino acid residues which are predominantly hydrophobic to form hydrophobic interactions with the lipid bilayer of the cell membrane. Such membrane proteins are also known as integral membrane proteins. In most cases, the integral membrane proteins extend through the cell membrane into the interior of the cell, thus comprising an extracellular domain, one or more transmembrane domains and an intracellular domain.

Many membrane proteins have become important targets for drug development in that irregularities in cellular signalling, that is, either increased or decreased signal transduction, from these proteins have been found to play a role in the development of a wide variety of diseases. Recent drug development has therefore focused on screening for compounds which are capable of either upregulating or downregulating the activity of the membrane protein, as required. Screening has usually been performed in a "shot-gun" fashion by setting up an assay for screening large numbers of compounds, e.g. compounds in combinatorial libraries, to identify compounds with the desired activity. Optimization of the hits from such screening procedures has been quite cumbersome and resource-demanding, involving procedures such as described by E. Sun and F.E. Cohen,

Gene 1993 137(1),127-32, or J. Kuhlmann, *Int J Clin Pharmacol Ther.* 1999 37(12), 575-83. A major disadvantage of the classical drug discovery process is that it is difficult to identify active compounds with sufficient s lectivity and specificity for a given target membrane protein and the desired pharmaco-kinetic properties.

For ligand optimization purposes, it is often an advantage to elucidate the ligand binding sites of various membrane proteins. The information obtained from these various procedures may be used to acquire a better understanding of ligand binding to membrane proteins involved in dysfunction or disease which is important in the design of new selective drugs towards the target in question.

Thus, site-directed mutagenesis is used to eliminate a ligand binding site or part of a ligand binding site by substitution of selected amino acid residues with other residues, e.g. alanine. Only a few cases have been presented where ligand binding sites have been thoroughly investigated by an extensive and systematic mutational analysis of all possible residues in a given area (e.g. the β-adrenergic receptor, Strader et al., *FASEB J. 3*, 1989, pp. 1825-1832; Strader et al, *J. Biol. Chem. 266*(1), 1991, pp. 5-8). A general problem of the site-directed mutagenesis method is that it is not clear whether the substitution of a residue affects the binding of a ligand directly (i.e. the residue is directly involved in ligand binding) or indirectly (i.e. the residue is only involved in the structure of the receptor).

Another problem of Ala substitution is false negative results because the procedure basically creates another "hole" in the presumed binding pocket through removal of the side chain on the residue replaced by Ala. The effect of Ala substitution is highly dependent on the relative contribution to the binding energy of the replaced residue. An alternative to Ala substitution is to introduce a larger side chain, e.g. Trp, in a presumed binding pocket as described by Holst et al., *Mol Pharmacol. 53*(1), 1998, pp. 166-175. This could cause many other problems for the ligand by impairing the interaction not only with the mutated residue but presumably also with neighbouring residues due to the incongruence that would be created in a larger part of the interface.

Exchange of all polar hydrogen-bond forming residues in the outer part of the receptor has been used to identify major interaction sites for non-peptide antagonists in the tachykinin receptors (Fong et al., *Nature 362*, 1993, pp. 350-353).

Conserved residues and non-conserved residues between homologous receptors may reveal useful information in the case where particular residues are of interest. For instanc , for investigation of binding interactions of a positiv ly charged ligand to acidic residues, all conserved Asp and Glu residues can be substituted with Asn or Gln in the event that both receptors bind the same ligand. On the other hand, all non-conserved Asp and Glu residues can be substituted in the event that only one of the receptors bind the ligand.

Chimeric receptor constructs can be made to locate domains involved in ligand binding. The method was first described for G-protein coupled receptors by Kobilka et al, *Science 240*, 1988, pp. 1310-1316. By combining a non-ligand binding receptor with a ligand-binding receptor in different ways, useful information can be obtained to pinpoint the area of binding. Most often chimeric receptors are constructed between receptors with a certain degree of homology (e.g. NK-1 and NK-2, Gether et al., *Proc. Natl. Acad. Sci. USA 90*(13), 1993, 6194-6198), or the same receptor from different species (e.g. hCXR4 and rCXR4, Labrosse et al., *J. Virol. 72*(8), 1998, 6381-6388).

Photoaffinity labelling has been proven to be a useful tool in identifying domains of receptors involved in ligand binding (Dohlman et al., *Ann. Rev. Biochem. 60*, 1991, pp. 653-688). A photoreactive group is attached or built into the ligand. After binding, the ligand-receptor complex is exposed to UV light, resulting in crosslinking of the ligand to the receptor. Finally the complex is digested with proteases and the ligand-binding part of the receptor can be identified.

Structure-activity relationships (SAR) can provide a great deal of information regarding the nature of ligand-receptor interactions, but no information about the location of the binding site is provided. A number of closely related chemical structures are used to direct the orientation of the ligand within the putative binding cavity and to determine what part of the ligand is involved in binding to the receptor. This technique has its limitations due to the fact that changing the structure of the ligand may result in a change in the binding site of the receptor. The information obtained from the SAR can be used to build a pharmacophore model which contains information as to the parts of the ligand that are important for the binding and activity of the ligand. SAR in combination with site-directed mutagenesis, e.g. testing different ligands on different receptor mutants, may reveal interesting structural information, e.g. distance constraints within the receptor.

Finally, determination of the crystal structure by X-ray crystallography of three-dimensional crystals provides very high resolution structures and consequently very high quality information about the various structural features of the membrane protein. However, crystal structures are very difficult and consequently very expensive to provide, and the crystal structure has therefore only been elucidated for a few of the vast number of membrane proteins of potential interest as drug discovery targets.

A wide spectrum of ligands have been identified for the various families of integral membrane proteins. The types of chemical messengers acting through membrane proteins include ions, amino acids, monoamines, lipids, purines, neuropeptides, peptide hormones, chemokines, glycoprotein hormones and proteases. It has been found that several integral membrane proteins include binding sites for metal ions. The coordination of metal ions to metal ion binding sites is well characterized in numerous high-resolution X-ray structures of soluble proteins; for example, distances from the chelating atoms to the metal ion as well as the preferred conformation of the chelating side chains are known (e.g. J.P. Glusker, *Adv. Protein Chem. 42*, 1991, pp. 3-76; P. Chakrabarty, *Protein Eng. 4*, 1990, pp. 57-63; R. Jerigan et al., *Curr. Opin. Struct. Biol. 4*, 1994, pp. 256-263). Hence, characterizing a metal ion binding site in a membrane protein using, for example, molecular models and site directed mutagenesis yields information about the structure of the membrane protein and importantly where the "ligand" (metal ion) binds (e.g. Elling et al. *Fold. Des.* 2(4), 1997, pp. S76-80).

Engineering artificial metal ion binding sites in membrane proteins has also been employed to explore the structure of the protein. If side chains of amino acid residues participating in metal ion binding are known, it imposes a distance constraint on the protein structure which can be used in the interpretation of unknown protein structures (C.E. Elling and T.W. Schwartz, *EMBO J. 15*(22), 1996, pp. 6213-6219; C.E. Elling et al., *Fold. Des. 2*(4), 1997, pp. S76-80). C.E. Elling et al., *Nature 374*, 1995, pp. 74-77, have reported the conversion of an antagonist binding site in the tachykinin NK-1 receptor to a metal ion binding site. The metal ion binding sites created so far have all been antagonistic, i.e. generating inactive membrane proteins, but recently the generation of an activating metal ion binding site has been reported for the β_2-adrenergic receptor (C.E. Elling et al, *PNAS 96*, 1999, pp. 12322-12327).

SUMMARY OF THE INVENTION

The present invention utilises the finding that it is possible to mutate proteins such as a receptor, an enzyme or a transcriptional regulator in such a way that they comprise a metal ion binding site. In this way, it is possible to construct test compounds which bind to such mutated proteins at a site in the protein determined by the location of the artificially engineered metal ion binding site but which are also optimised for interaction with spatially neighbouring parts of the proteins (that is, interaction with side chains of one or more neigbouring amino acid residues). These compounds can then be utilized as chemical scaffolds or starting points for the construction of compounds binding to the wild-type protein. In this way it is possible to predetermine the binding site of a compound to a particular location in a protein structure and thereby target the optimised compounds to sites where binding of the compound will alter the biological activity of the protein in a desired way, for example to increase or decrease its biological activity. By selecting the binding site for a test compound at will and thereby select the binding site for the optimised compound (such as a drug candidate) in a protein, it is for example 1) possible to speed up the process of development of high affinity drug candidates or other compounds with biological activity because a more efficient structure-based compound optimisation process can be applied; 2) possible to obtain high selectivity for a given member of a protein family by targeting the compound to a site in the protein which differs between different members of the protein family; 3) possible to obtain new functionalities of compounds by targeting them to allosteric modulatory sites in proteins.

In the course of research leading to the present invention, the inventors have found that certain small organic molecules which bind metal ions (i.e. metal ion chelators) are also able to bind to metal ion binding sites in various proteins, including membrane proteins, such as receptors, in such a way that the metal ion acts as a bridge between the small organic molecule and the protein. These compounds bind with affinities corresponding to the binding affinities of most lead compounds discovered by simple screening of large chemical libraries. Importantly, the present invention has made it possible to identify and localise the exact binding site and binding mode of such metal ion chelates used as test compounds, contrary to what has generally been known in the art about the binding of test compounds or even high-affinity drugs. Based on the identification or confirmation of the binding site of the test compounds, using for example site-directed mutagenesis, three-dimensional structure or molecular models of the protein and techniques such as those

described abov , a rational approach may be taken to the chemical optimisation of th test compounds. Thus, relatively small chemical libraries may be made, the compounds in which may be designed to interact with specific amino acid residues of th prot in in question. Compounds that exhibit a high affinity binding to the protein and affect the biological activity of the protein in a desired way may then be selected for further optimisation.

The metal ion portion of the metal ion chelates used as test compounds may subsequently be removed, and the test compounds, as well as chemical derivatives thereof constructed to interact with side chains of other amino acids in the vicinity of the artificial metal ion binding site, may then be tested for binding to the wild-type protein which does not include a metal ion binding site. Relatively small chemical libraries may be made, the compounds in which may be designed to interact with the specific amino acid residues found in the wild-type protein at the location where the metal ion site had initially been engineered.

Thus, the present invention is based on the general principle, applicable to any protein, of introducing metal ion binding sites at any position in the protein where a test compound binding to the protein is likely to exert an effect on the biological activity of the protein. This may for example be 1) at a site where the protein binds to another protein, for example a regulatory protein; 2) at a site where the binding of the test compound will interfere with the cellular targeting of the protein; 3) at a site where the binding of the test compound will directly or indirectly interfere with the binding of substrate or the binding of an allosteric modulatory factor for the protein; 4) at a site where the binding of the test compound may interfere with the intra-molecular interaction of domains within the protein, for example the interaction of a regulatory domain with a catalytic domain; 5) at a site where binding of the test compound will interfere with the folding of the protein, for example the folding of the protein into its active conformation; or 6) at a site which will interfere with the activity of the protein, for example by an allosteric mechanism. Subsequent to identifying test compounds that bind to the artificial metal ion binding site of the protein, information may be acquired of the structure of the binding site and of amino acid residues in its immediate vicinity. Such information may be used in the design of compounds with improved binding affinity to the proteins resulting from interaction with one or more amino acid residues in the vicinity of the metal ion binding site. Such compounds may, in turn, be used in

the design of potential drug candidates or other compounds with a desired activity on the corresponding wild-type, non-mutated protein.

Accordingly, the present invention relates to a method of identifying a compound capable of interacting with a protein so as to modify its biological activity, the method comprising

(a) mutating a protein by substituting at least one amino acid residue by at least one other amino acid residue capable of binding a metal ion, resulting in the introduction of a metal ion binding site into the protein;
(b) contacting the mutated protein with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the protein, and determining any effect of the binding of said test compound to said protein by detecting any change in the biological activity of said protein; and
(c) contacting said protein in wild-type, non-mutated form with at least one test compound determined to bind the mutated protein in step (b) but lacking a metal ion chelated thereto, and determining any interaction of the test compound lacking the metal ion with the wild-type protein.

A very important class of proteins amenable to testing according to the present invention are membrane proteins which includes proteins that are involved in intercellular communication and other biological processes of profound importance for cellular activity. Thus, in another aspect, the present invention relates to a method of identifying compounds acting as ligands of a membrane protein, the method comprising

(a) mutating a membrane protein or putative membrane protein by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the membrane protein;
(b) contacting the mutated membrane protein with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the membrane protein, and determining any binding of said test compound to said membrane protein; and

(c) contacting said membrane protein in wild-type non-mutated form, with at least one test compound determined to bind to the mutated membrane protein in step (b) but lacking a metal ion chelated thereto, and determining any interaction of the test compound lacking the metal ion with the wild-type membrane protein.

In a further aspect, the invention relates to a method of identifying a compound acting as a ligand of an orphan receptor, the method comprising

(a) mutating an orphan receptor or putative orphan receptor by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the orphan receptor;
(b) contacting the mutated orphan receptor with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the orphan receptor, and determining any binding of said test compound to said orphan receptor.

In a still further aspect, the invention relates to a method of screening for compounds capable of acting as ligands on an orphan receptor, the method comprising contacting an orphan receptor or putative orphan receptor which is mutated by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the orphan receptor, with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue capable of binding metal ions, as well as one or more test substances which are not metal ion chelates, under conditions permitting binding of the metal ion chelate test compound and the one or more test substances to the mutated orphan receptor, and determining any binding of said test substance or substances to said mutated orphan receptor by measuring altered binding of the metal chelate test compound to the receptor or changes in signal transduction from the orphan receptor.

DETAILED DESCRIPTION OF THE INVENTION

The following d tailed description of the invention is mainly concerned with methods of identifying compounds interacting with membrane proteins. It should be understood, however, that the discussion of the detailed method steps apply equally to proteins in general.

In the following description and claims, the following terms shall be defined as indicated below.

A "test compound" is intended to indicate any drug, substance, compound or molecule with potential biological activity. The term includes in its meaning metal ion chelates of the general formulae shown below as well as chemical derivatives thereof constructed to interact with side chains of other amino acids in the vicinity of the artificial metal ion binding site.

A "metal ion chelator" is intended to indicate a compound capable of forming a complex with a metal ion. Such a compound will generally contain a heteroatom such as N, O or S with which the metal ion is capable of forming a complex.

A "metal ion chelate" is intended to indicate a complex of a metal ion chelator and a metal ion.

A "metal ion binding site" is intended to indicate an amino acid residue of a protein which comprises an atom capable of complexing with a metal ion. Such an atom will typically be a heteroatom, in particular N, O or S.

A "ligand" is intended to include any substance that either inhibits or stimulates the activity of the membrane protein or that competes for the receptor in a binding assay. An "agonist" is defined as a ligand increasing the functional activity of a membrane protein (e.g. signal transduction through a receptor). An "antagonist" is defined as a ligand decreasing the functional activity of a membrane protein either by inhibiting the action of an agonist or by its own intrinsic activity. An "inverse agonist" (also termed "negative antagonist") is defined as a ligand decreasing the basal functional activity of a membrane protein.

A "protein" is intended to include any protein with a discernible biological activity in any unicellular or multicellular organism, including bacteria, fungi, plants, insects, animals or mammals, including humans. Thus, the protein may suitably be a drug target, i.e. a protein whose activity is important for the development or amelioration of a disease state, and whose level of activity may be altered (i.e. up- or downregulated) due to the influence of a biologically active substance such as a drug, for example a compound with abortifacient, acromegaly, alcohol deterrent, amebicide, anabolic, analeptic, analgesic, anesthetic, antiacne, antiallergic, antiallergic, ophthalmic, anti-Alzheimer's disease, antianginal, antiarrhythmic, antiarthritic, antiasthmatic, antibacterial, antibiotic, anticancer, anticholelithogenic, anticoagulant, anticonvulsant, antidepressant, antidiabetic, antidiarrheal, antiemetic, antiepileptic, antiestrogen, antifungal, antiglaucoma, antihistamine, antihypertensive, antiinflammatory, antilipidemic, antimalarial, antimigraine, antinauseant, antineoplastic, antiobesity, antiparasitic, antiparkinsonian, antiperistaltic, antiprogestogen, antiprolactin, antiprostatic hypertrophy, antipsoriatic, antipsychotic, antirheumatic, antisecretory, antiseptic, antispasmodic, antithrombotic, antitussive, antiulcer, antiviral, anxiolytic, bronchodilator, calcium regulator, cardioprotective, cardiostimulant, cardiotonic, cephalosporin, cerebral vasodilator, chelator, choleretic chrysotherapeutic, cognition enhancer, congestive heart failure, coronary vasodilator, cystic fibrosis, cytoprotective, dependence treatment, diuretic, dyslipidemia, enzyme, expectorant, fertility enhancer, fibrinolytic, gastroprokinetic, gaucher's disease, growth hormone, growth hormone insensitivity, haemophilia, heart failure, hematologic, hematopoetic, hemostatic, hepatroprotective, hormone, hyperphenylalaninemia, hyperprolactinemia, hypertensive, hypnotic, hypoammonnuric, hypocalciuric, hypocholesterolemic, hypoglycemia, hypolipaemic, hypolipidemic, idiopathic hypersomnia, immunomodulator, immunostimulant, immunosuppressant, beta-lactamase inhibitor, leukopenia, lung surfactant, mucolytic, muscle relaxant, multiple sclerosis, muscle relaxant, narcotic antagonist, nasal decongestant, neuroleptic, neuromuscular blocker, neuroprotective blocker, neuroprotective, nootropic, non-steroid antiinflammatory disease (NSAID), osteoporosis, paget's disease, platelet aggregation inhibitor, platelet antiaggregant, pneumonia, precocious puberty progestogen, protease inhibitor, psychostimulant, 5-alpha-reductase inhibitor, respiratory surfactant, subarachnoid hemorrhage, thrombolytic, ulcerative colitis, urolithiasis, urologic, vasoprotective, vulnerary and wound healing properties. Important proteins for the present purpose are proteins which may be stabilised in an active or inactive conformation by a biologically active substance. In this way, it may be possible to obtain an effect of a test compound of the type described herein irrespective of whether the active site of the protein is known, or whether the

structure of the activ site has been r solved (e.g. by X-ray crystallisation) Examples of such proteins are enzymes, receptors, hormones and other signalling molecules, transcriptional factors and regulators, intra- or extracellular structural proteins, in particular actins; adaptins; antibodies; ATPases; cyclins; dehydrogenases; GTP-binding proteins; GTP/GDP-exchange factors; GTPase activating proteins; GTP/GDP dissociation inhibitors; chaperones; histones; histone acetyltransferases & deacetyltransferases; hormones and other signalling proteins and peptides; kinases; lipases; major facilitator superfamily proteins; motorproteins; nucleases; polymerases; isomerases; proteases; protease inhibitors; phosphatases; ubiquitin-system proteins; membrane proteins including receptors, transporters and channels; transcription factors and tubulins; preferably membrane receptors; nuclear receptors, zinc finger proteins; proteases, tyrosine kinases and matrix proteins. Other important proteins for the present purpose are proteins whose biological activity is regulated by their cellular targeting and whose biological activity therefore can be modulated by drugs which alter their cellular targeting with or without altering their actual intrinsic activity.

A "membrane protein" is intended to include any protein anchored in a cell membrane and mediating cellular signalling from the cell exterior to the cell interior. Important classes of membrane proteins include receptors such as tyrosine kinase receptors, G-protein coupled receptors, adhesion molecules, ligand- or voltage-gated ion channels, or enzymes. The term is intended to include membrane proteins whose function is not known, such as orphan receptors. In recent years, largely as part of the human genome project, large numbers of receptor-like proteins have been cloned and sequenced, but their function is as yet not known. The present invention may be of use in elucidating the function of the presumed receptor proteins by making it possible to develop methods of identifying ligand for orphan receptors based on compounds developed from metal ion chelates that bind to mutated orphan receptors into which artificial metal ion binding sites have been introduced.

"Signal transduction" is defined as the process by which extracellular information is communicated to a cell by a pathway initiated by binding of a ligand to a membrane protein, leading to a series of conformational changes resulting in a physiological change in the cell in the form of a cellular signal.

A "functional group" is intended to indicate a chemical ntity which is a component part of the test compound and which is capable of interacting with an amino acid residue or a side chain of an amino acid residue of the membrane protein. Examples of such functional groups include, but are not limited to, electrophiles such as the ammonium ion, or hydrogen bond donor or acceptor groups such as amino, amide, carboxy, sulphonate, etc.

A "wild-type" membrane protein is understood to be a membrane protein in its native, non-mutated form, in this case not comprising an introduced metal ion binding site.

The method of the invention typically includes the following steps:

(a) subjecting a nucleotide sequence encoding, or believed to encode, a membrane protein to mutagenesis so as to substitute at least one amino acid residue by another amino acid residue capable of binding a metal ion,
(b) expressing said nucleotide sequence in a suitable host cell, or
(c) contacting said cell or a portion thereof including the expressed membrane protein with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to at least one amino acid residue of said membrane protein, under conditions permitting binding of the test compound to the membrane protein, and determining any binding of said test compound to said membrane protein.

As used herein the term "nucleotide sequence" is intended to indicate any nucleic acid molecule of cDNA, genomic DNA, synthetic DNA or RNA origin. The nucleotide sequence may be single- or double-stranded, and may be based on a complete or partial naturally occurring nucleotide sequence encoding a membrane protein of interest to the present invention. The nucleotide sequence may optionally contain other nucleic acid segments.

The nucleotide sequence may suitably be of genomic or cDNA origin, for instance obtained by preparing a genomic or cDNA library and screening for DNA sequences coding for all or part of the polypeptide by hybridization using synthetic oligonucleotide probes in accordance with standard techniques (cf. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2. Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989). For the present purpose, the DNA sequence encoding the polypeptide is preferably of

vertebrate origin, .g. derived from a mammalian, preferably human genomic DNA or cDNA library.

The nucleotide sequence encoding the membrane protein may also be prepared synthetically by established standard methods, e.g. the phosphoamidite method described by Beaucage and Caruthers, *Tetrahedron Letters 22* (1981), 1859-1869, or the method described by Matthes et al., *EMBO Journal 3* (1984), 801-805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned in suitable vectors.

Furthermore, the nucleotide sequence may be of mixed synthetic and genomic, mixed synthetic and cDNA or mixed genomic and cDNA origin prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate), the fragments corresponding to various parts of the entire nucleotide sequence, in accordance with standard techniques.

The nucleotide sequence may also be prepared by polymerase chain reaction using specific primers, for instance as described in US 4,683,202 or Saiki et al., *Science 239* (1988), 487-491.

In step (a) of the present method, the nucleotide sequence encoding the membrane protein may be subjected to site-directed mutagenesis in order to introduce the amino acid residue which includes the metal ion binding site. Site-directed mutagenesis may be performed according to well-known techniques, e.g. as described in Ho et al., *Gene 77*, 51-59; Horton et al., Gene, 77, 61-68 or *"Current protocols in Molecular Biology"*, Ausubel et al., John Wiley & Sons Inc., ISBN 0-471-50338-X & 0-471-50337-1. In order to generate a metal ion binding site in the membrane protein, one or more amino acid residues capable of binding metal ions are introduced. Amino acid residues which function as effective metal ion binding residues are typically those that contain electron donating atoms such as S, O or N. Alternatively, metal ion binding amino acid residues may be found among aromatic amino acids. Thus, preferred candidate amino acids for binding metal ions, and hence for introduction into the membrane protein, are those selected from Ser, Lys, Arg, Tyr, Thr, Trp, Phe, Asp, Glu, Asn, Gln, Cys and His, in particular Asp, Glu, Cys and His, preferably His.

The nucleotide sequence encoding the membrane protein may suitably be inserted into a suitable vector which may be any vector which may conveniently be subjected to recombinant DNA procedures. The choice of vector will often depend on the host cell into which it is to be introduced. Thus, the vector may be an autonomously replicating vector, i.e. a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g. a plasmid. Alternatively, the vector may be one which, when introduced into a host cell, is integrated into the host cell genome and replicated together with the chromosome(s) into which it has been integrated.

The vector is preferably an expression vector in which the nucleotide sequence encoding the membrane protein is operably linked to additional segments required for transcription of the DNA. In general, the expression vector is derived from plasmid or viral DNA, or may contain elements of both. The term "operably linked" indicates that the segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in a promoter and proceeds through the nucleotide sequence coding for the membrane protein. The promoter may be any DNA sequence which shows transcriptional activity in the host cell of choice and may be derived from genes encoding proteins either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleotide sequence encoding the polypeptide of the invention in mammalian cells are the ubiquitin promoter (Wiborg et al., *EMBO J. 4* (1985), 755-759), the SV40 promoter (Subramani et al., *Mol. Cell Biol. 1* (1981), 854-864), the MT-1 (metallothionein gene) promoter (Palmiter et al., *Science 222* (1983), 809-814) or the adenovirus 2 major late promoter.

An example of a suitable promoter for use in insect cells is the polyhedrin promoter (US 4,745,051; Vasuvedan et al., *FEBS Lett. 311*, (1992) 7-11), the P10 promoter (J.M. Vlak et al., *J. Gen. Virology 69*, 1988, pp. 765-776), the Autographa californica polyhedrosis virus basic protein promoter (EP 397 485), the baculovirus immediate early gene 1 promoter (US 5,155,037; US 5,162,222), or the baculovirus 39K delayed-early gene promoter (US 5,155,037; US 5,162,222).

Examples of suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., *J. Biol. Chem. 255* (1980), 12073-12080; Alber and Kawasaki, *J. Mol. Appl. Gen. 1* (1982), 419-434) or alcohol dehydrogenase genes (Young et

al., in *Genetic Engineering of Microorganisms for Chemicals* (Hollaender et al., eds.), Plenum Press, New York, 1982), or the TPI1 (US 4,599,311) or ADH2-4c (Russell et al., *Nature 304* (1983), 652-654) promoters.

The nucleotide sequence encoding the membrane protein may also, if necessary, be operably connected to a suitable terminator, such as the human growth hormone terminator Palmiter et al., op.cit.) or (for fungal hosts) the TPI1 (Alber and Kawasaki, op. cit.) or ADH3 (McKnight et al., op. cit.) terminators. The vector may further comprise elements such as polyadenylation signals (e.g. from SV40 or the adenovirus 5 Elb region), transcriptional enhancer sequences (e.g. the SV40 enhancer) and translational enhancer sequences (e.g. the ones encoding adenovirus VA RNAs).

The recombinant vector may further comprise a DNA sequence enabling the vector to replicate in the host cell in question. An example of such a sequence (when the host cell is a mammalian cell) is the SV40 origin of replication.

When the host cell is a yeast cell, suitable sequences enabling the vector to replicate are the yeast plasmid 2μ replication genes REP 1-3 and origin of replication.

The vector may also comprise a selectable marker, e.g. a gene the product of which complements a defect in the host cell, such as the gene coding for dihydrofolate reductase (DHFR) or the Schizosaccharomyces pombe TPI gene (described by P.R. Russell, *Gene 40*, 1985, pp. 125-130), or one which confers resistance to a drug, e.g. ampicillin, kanamycin, tetracyclin, chloramphenicol, neomycin, hygromycin or methotrexate.

The procedures used to ligate the nucleotide sequences coding for the membrane protein, the promoter and optionally the terminator sequence, respectively, and to insert them into suitable vectors containing the information necessary for replication, are well known to persons skilled in the art (cf., for instance, Sambrook et al., op.cit.).

The nucleotide sequence encoding the membrane protein introduced into the host cell may be either homologous or heterologous to the host in question. If homologous to the host cell, i.e. produced by the host cell in nature, it will typically be operably connected to another promoter sequence or, if applicable, another terminator sequence than in its natural environment. The term "homologous" is intended to include a nucleotide sequence en-

coding a polypeptide native to the host organism in question. The term "heterologous" is intended to include a nucleotide sequence not expressed by the host cell in nature. Thus, the nucleotide sequence may be from another organism, or it may be a synthetic sequence.

The host cell into which the nucleotide sequence or the recombinant vector is introduced may be any cell which is capable of producing the present membrane protein and includes yeast cells and higher eukaryotic cells such as insect or mammalian cells.

Examples of suitable mammalian cell lines are the COS (ATCC CRL 1650 and 1651), BHK (ATCC CRL 1632, ATCC CCL 10), CHL (ATCC CCL39), CHO (ATCC CCL 61), HEK293 (ATCC CRL 1573) or NIH/3T3 (ATCC CRL 1658) cell lines. Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in e.g. Kaufman and Sharp, *J. Mol. Biol. 159* (1982), 601-621; Southern and Berg, *J. Mol. Appl. Genet. 1* (1982), 327-341; Loyter et al., *Proc. Natl. Acad. Sci. USA 79* (1982), 422-426; Wigler et al., *Cell 14* (1978), 725; Corsaro and Pearson, *Somatic Cell Genetics 7* (1981), 603, Graham and van der Eb, *Virology 52* (1973), 456, Neumann et al., *EMBO J.* (1982), 841-845; and in accordance with standard techniques (cf. Celis et al., *Cell Biology: A Laboratory Handbook*, 2. Ed. Academic Press, ISBN 0-12-164726-9; and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2. Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989).

Examples of suitable yeast cells include cells of Saccharomyces spp. or Schizosaccharomyces spp., in particular strains of Saccharomyces cerevisiae or Saccharomyces kluyveri. Methods for transforming yeast cells with heterologous DNA and producing heterologous polypeptides therefrom are described, e.g. in US 4,599,311, US 4,931,373, US 4,870,008, 5,037,743, and US 4,845,075, all of which are hereby incorporated by reference. Transformed cells are selected by a phenotype determined by a selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient, e.g. leucine. A preferred vector for use in yeast is the POT1 vector disclosed in US 4,931,373.

The DNA sequence encoding the polypeptide of the invention may be preceded by a signal sequence and optionally a leader sequence, e.g. as described above. Further examples of suitable yeast cells are strains of Kluyveromyces, such as K. lactis, Hansenula,

.g. H. polymorpha, or Pichia, e.g. P. pastoris (cf. Gleeson et al., *J. Gen. Microbiol. 132*, 1986, pp. 3459-3465; US 4,882,279).

Transformation of insect cells and expression of heterologous membrane protein therein may be performed as described in US 4,745,051; US 4,879,236; US 5,155,037; 5,162,222; EP 397,485, all of which are incorporated herein by reference. The insect cell line used as the host may suitably be a Lepidoptera cell line, such as Spodoptera frugiperda cells or Trichoplusia ni cells (cf. US 5,077,214). Culture conditions may suitably be as described in, for instance, WO 89/01029 or WO 89/01028, or any of the aforementioned references.

The transformed or transfected host cell described above is then cultured in a suitable nutrient medium under conditions permitting the expression of the membrane protein, after which the cell expressing the membrane protein is recovered from the culture and used in a suitable assay to determine the activity of the membrane protein in the presence of test compounds.

The medium used to culture the cells may be any conventional medium suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection).

The membrane protein used in the method of the invention is suitably an integral membrane protein, which is to say a membrane protein anchored in the cell membrane. The membrane protein is preferably of a type comprising at least one transmembrane domain. Interesting membrane proteins for the present purpose are found in classes comprising 1-14 transmembrane domains.

Thus, membrane proteins of interest comprising one transmembrane domain are receptors such as tyrosine kinase receptors, e.g. a growth factor receptor such as the growth hormone, insulin, epidermal growth factor, transforming growth factor, erythropoietin, colony-stimulating factor, platelet-derived growth factor receptor or nerve growth factor receptor (TrkA or TrkB).

Membrane proteins of interest comprising two transmembrane domains are, .g., purinergic ion channels.

Membrane proteins of interest comprising 3, 4 or 5 transmembrane domains are, e.g., ligand-gated ion channels, such as nicotinic acetylcholine receptors, GABA receptors, or glutamate receptors (NMDA or AMPA).

Membrane proteins of interest comprising 6 transmembrane domains are, e.g., voltage-gated ion channels, such as potassium, sodium, chloride or calcium channels.

Membrane proteins of interest comprising 7 transmembrane domains are, e.g., G-protein coupled receptors, such as the acetylcholine, adenosine, adrenoceptors, anaphylatoxin chemotactic receptor, angiotensin, bombesin (neuromedin), bradykinin, calcitonin, calcitonin gene related peptide, conopressin, corticotropin releasing factor, amylin, adrenomedullin, calcium, cannabinoid, chemokine, cholecystokinin, conopressin, corticotropin-releasing factor, dopamine, eicosanoid, endothelin, fMLP, $GABA_B$, galanin, gastrin, gastric inhibitory peptide, glucagon, glucagon-like, glutamate, glycoprotein hormone (e.g. FSH, LSH, TSH, LH), gonadotropin releasing hormone, growth hormone releasing hormone, growth hormone secretagogue, histamine, 5-hydroxytryptamine, leukotriene, lysophospholipid, melanocortin, melatonin, neuropeptide Y, neurotensin, olfactory, opioid, opioid-like, opsin, orexin, oxytocin, parathyroid hormone/parathyroid hormone-related, P2Y, pheromone, platelet-activating factor, prostanoid, protease-activated, secretin, somatostatin, tachykinin, thrombin and protease activated, thyrotropin-releasing hormone, pituitary adenylate activating peptide, vasopressin, vasoactive intestinal peptide and viral encoded receptors; in particular galanin, P2Y, chemokine, melanocortin, bombesin, cannabinoid, lysophospholipid, fMLP, neuropeptide Y, tachykinin, dopamine, histamine, 5-hydroxytryptamine, histamine, mas-proto-oncogene, acetylcholine, oxytocin, herpes virus encoded, epstein-barr virus encoded, cytomegalovirus encoded and bradykinin receptors; preferably galanin receptor type 1, leukotriene B4 receptor, CXCR5, melanocortin-1 receptor, melanocortin-3 receptor, melanocortin-4 receptor, melanocortin-5 receptor, bombesin receptor subtype 3, cannabinoid receptor 1, cannabinoid receptor 2, EDG-2, EDG-4, FMLP-related receptor I, FMLP-related receptor II, Y6 receptor, Y5 receptor, NK-1 receptor, NK-3 receptor, D2 receptor (short), D2 receptor (long), D4 receptor (D2C), duffy antigen, Q89609 herpes virus type 2, VU51_HSV6U, histamine H1 receptor, MAS proto-oncogene, muscarinic M1 receptor, muscarinic M2 receptor, muscarinic M3 receptor,

muscarinic M5 receptor, oxytocin receptor, XCR1 receptor, EBI2 receptor, RDC1 receptor, O76000 r ceptor, O15218 r ceptor, GPR12 receptor or GPR3 receptor.

Membrane proteins of interest comprising 8, 10, 12 or 14 transmembrane domains are, e.g., transporter proteins, such as a GABA, monoamine or nucleoside transporter.

The membrane protein may also be a multidrug resistance protein, e.g. a P-glycoprotein, multidrug resistance associated protein, drug resistance associated protein, lung resistance related protein, breast cancer resistance protein, adenosine triphosphate-binding cassette protein, Bmr, QacA or EmrAB/TolC pump.

The membrane protein may also be a cell adhesion molecule, e.g. NCAM, VCAM or ICAM.

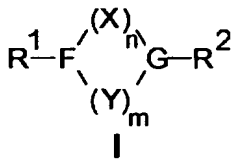
Furthermore, the membrane protein may be an enzyme such as adenylyl cyclase.

As indicated above, the present invention also relates to a method of identifying compounds such as for example abortifacient, acromegaly, alcohol deterrent, anti-Alzheimer's disease, amebicide, anabolic, analeptic, analgesic, anesthetic, antiacne, antiallergic, antiallergic, ophthalmic, antianginal, antiarrhythmic, antiarthritic, antiasthmatic, antibacterial, antibiotic, anticancer, anticholelithogenic, anticoagulant, anticonvulsant, antidepressant, antidiabetic, antidiarrheal, antiemetic, antiepileptic, antiestrogen, antifungal, antiglaucoma, antihistamine, antihypertensive, antiinflammatory, antilipidemic, antimalarial, antimigraine, antinauseant, antineoplastic, antiobesity, antiparasitic, antiparkinsonian, antiperistaltic, antiprogestogen, antiprolactin, antiprostatic hypertrophy, antipsoriatic, antipsychotic, antirheumatic, antisecretory, antiseptic, antispasmodic, antithrombotic, antitussive, antiulcer, antiviral, anxiolytic, bronchodilator, calcium regulator, cardioprotective, cardiostimulant, cardiotonic, cephalosporin, cerebral vasodilator, chelator, choleretic chrysotherapeutic, cognition enhancer, congestive heart failure, coronary vasodilator, cystic fibrosis, cytoprotective, dependence treatment, diuretic, dyslipidemia, enzyme, expectorant, fertility enhancer, fibrinolytic, gastroprokinetic, gaucher's disease, growth hormone, growth hormone insensitivity, haemophilia, heart failure, hematologic, hematopoetic, hemostatic, hepatroprotective, hormone, hyperphenylalaninemia, hyperprolactinemia, hypertensive, hypnotic, hypoammonnuric, hypocalciuric, hypocholesterolemic, hypoglycemia, hypolipaemic, hypolipidemic, idiopathic hypersomnia, immunomodulator, immunostimulant,

immunosuppressant, beta-lactamase inhibitor, leukopenia, lung surfactant, mucolytic, muscle relaxant, multiple sclerosis, muscle relaxant, narcotic antagonist, nasal decongestant, neuroleptic, neuromuscular blocker, neuroprotectiv blocker, neuroprotective, nootropic, non-steroid antiinflammatory disease (NSAID), osteoporosis, paget's disease, platelet aggregation inhibitor, platelet antiaggregant, pneumonia, precocious puberty progestogen, protease inhibitor, psychostimulant, 5-alpha-reductase inhibitor, respiratory surfactant, subarachnoid hemorrhage, thrombolytic, ulcerative colitis, urolithiasis, urologic, vasoprotective, vulnerary and wound healing compounds interacting with proteins other than membrane proteins. Examples of such proteins are enzymes, receptors, hormones and other signalling molecules, transcriptional factors and regulators, intra- or extracellular structural proteins in particular actins; adaptins; antibodies; ATPases; cyclins; dehydrogenases; GTP-binding proteins; GTP/GDP-exchange factors; GTPase activating proteins; GTP/GDP dissociation inhibitors; chaperones; histones; histone acetyltransferases & deacetyltransferases; hormones and other signalling proteins and peptides; kinases; lipases; major facilitator superfamily proteins; motorproteins; nucleases; polymerases; isomerases; proteases; protease inhibitors; phosphatases; ubiquitin-system proteins; transcription factors and tubulins; preferably membrane receptors; nuclear receptors; zinc finger proteins; proteases, tyrosine kinases and matrix proteins.

Test compounds which have been found to be useful in the present methods are typically compounds comprising a heteroalkyl or heterocyclyl moiety for chelating the metal ion. The term "heteroalkyl" is understood to indicate a branched or straight-chain chemical entity of 1-15 carbon atoms containing a heteroatom. The term "heterocyclyl" is intended to indicate a cyclic unsaturated, aromatic ("heteroaryl") or saturated ("heterocycloalkyl") group comprising at least one heteroatom. Preferred "heterocyclyl" groups comprise 5- or 6-membered rings with 1-3 heteroatoms or fused 5- or 6-membered rings comprising 1-4 heteroatoms. The heteroatom is typically N, O or S. Examples of heteroaryl groups are indolyl, dihydroindolyl, thiophenyl, furanyl, benzofuranyl, pyridinyl, pyrimidinyl, quinolinyl, triazolyl, imidazolyl, thiazolyl, tetrazolyl and benzimidazolyl. The heterocycloalkyl group generally includes 3-20 carbon atoms, and 1-4 heteroatoms.

Particularly useful test compounds are those of general formula I



I

wherein F is N, O or S and G is N, O or S;

and R^1 is an alkyl, aryl, cycloalkyl, alkoxy, ester, heteroalkyl, heterocycloalkyl or heteroaryl group, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide;

R^2 is an alkyl, aryl, cycloalkyl, alkoxy, ester, heteroalkyl or heteroaryl group, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide;

R^1 and/or R^2 optionally forming a fused ring together with the heteroatom to which it is attached, or R^1 and/or R^2 optionally forming a fused ring together with the heteroatom to which it is attached and a carbon atom of X or Y;

X is $-CH_2-$, CH_2-CH_2-, $-CH_2-S-CH_2-$, $-CH_2-N-CH_2-$, $-CH=CH-CH=CH-$, $-CH_2-(Z)_a-(P)_c-(W)_b-CH_2-$, $-CH_2-O-CH_2-$, wherein

each of Z and W are independently C, S, O or N, and

P is -CH- or $-CH_2-$;

Y is $-CH_2-$, CH_2-CH_2-, $-CH_2-S-CH_2-$, $-CH_2-N-CH_2-$, $-CH=CH-CH=CH-$, $-CH_2-(Z)_a-(P)_c-(W)_b-CH_2-$, $-CH_2-O-CH_2-$, wherein

each of Z and W are independently C, S, O or N, and

P is -CH- or $-CH_2-$;

n is 0 or an integer of 1-5,

m is 0 or an integer of 1-5,

a is an integer of 1-3,

b is an integer of 1-3, and

c is an integer of 1-7.

In the present context, the term "alkyl" is intended to indicate a branched or straight-chain, saturated or unsaturated chemical group containing 1-10, preferably 1-8, in particular 1-6 carbon atoms, such as methyl, ethyl, propyl, isopropyl, sec. butyl, tert. butyl, pentyl, isopentyl, hexyl, isohexyl, heptyl etc.

The term "cycloalkyl" is intended to denote a cyclic, saturated alkyl group of 3-7 carbon atoms.

The term "aryl" is intended to denote an aromatic (unsaturated), typically 5- or 6-membered, ring which may be a single ring (e.g. phenyl) or fused with other 5- or 6-membered rings (e.g. naphthyl or anthracyl).

The term "alkoxy" is intended to indicate the group alkyl-O-.

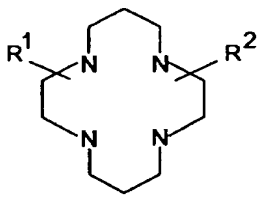
The term "amino" is intended to indicate the group -NR'R" where R' and R" are independently hydrogen, alkyl substituted alkyl, aryl, arylalkyl, substituted arylalkyl, heteroaryl, and substituted heteroaryl. In a primary amine group, both R' and R" are hydrogen, whereas in a secondary amino group, either but not both R' and R" is hydrogen.

The term "carboxy" is intended to indicate the group R'-COOH, where R' is as indicated above except hydrogen.

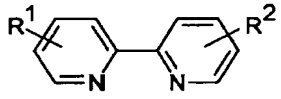
The term "ester" is intended to indicate the group COO-R', where R' is as indicated above except hydrogen.

Examples of halogen include fluorine, chlorine, bromine and iodine.

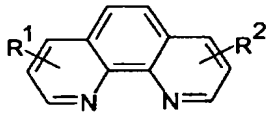
For the purpose of locating the presence of metal ion binding sites in proteins, test compounds in which the heterocyclic moiety is selected from a compound of formula IIa, IIb or IIc



IIa



IIb

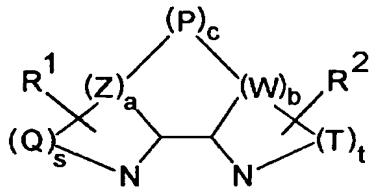


IIc

wherein R^1 and R^2 are as indicated above in formula I,

have been found to be particularly suitable.

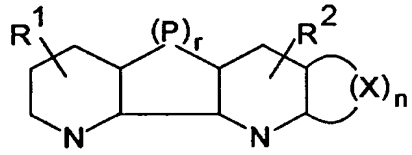
For the purpose of mapping the metal ion binding site or for scr ening for compounds capable of binding to the metal ion binding site of the protein, test compounds comprising a heterocyclic moiety of the general formula III



III

wherein R^1, R^2, Z, W, P, a, b and c are as indicated above, and
each of Q and T is independently -CH- or $-CH_2-$, s is an integer of 1-7, and t is an integer of 1-7, are believed to be particularly suitable.

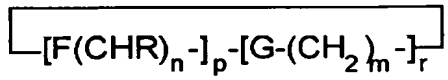
In particular, test compounds in which the heterocyclyl moiety has the general formula IV



IV

wherein R^1, R^2, P, X and n are as indicated above, and r is an integer of 1-3, are believed to be useful for mapping and screening purposes.

More preferred test compounds are those in which the heterocyclyl moiety has the general formula V

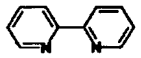


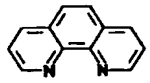
V

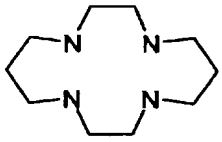
wherein F is N,O or S and G is N,O or S,
n is an integer from 1 to 5,

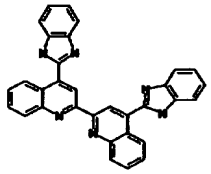
m is 0 or an integer from 1 to 5,
p is 0 or an integer from 1 to 8,
r is 0 or an integer from 1 to 8, and
R is alkyl, aryl, cycloalkyl, alkoxy, heteroalkyl or heteroaryl, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide.

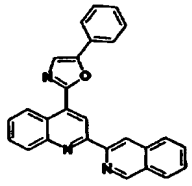
Examples of test compounds may be those in which the heterocyclic moiety is selected from a compound shown in Table 1:

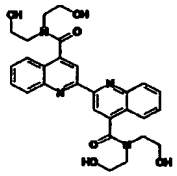
VII 

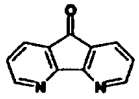
VIII 

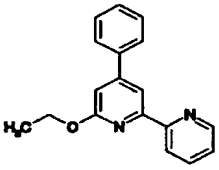
IX 

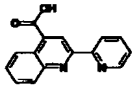
X 

XI 

XII 

XIII 

XIV 

XV 

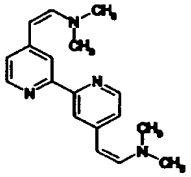
XVI 

Table I

Metal ions forming the complex with the heteroalkyl or heterocyclyl moiety in the test compounds may advantageously be selected from metal ions which have been tested for or are used for pharmaceutical purposes. Thus the metal ion is selected from the group consisting of rubidium, caesium, beryllium, magnesium, calcium, strontium, barium, scandium, yttrium, lutetium, titanium, zirconium, hafnium, vanadium, niobium, tantalum, chromium, molybdenum, tungsten, manganese, technetium, rhenium, iron, ruthenium, os-

mium, cobalt, rhodium, iridium, nick l, palladium, platinum, copper, silver, gold, zinc, cadmium, mercury, boron, aluminium, gallium, indium, thallium, silicon, germanium, tin, lead, arsenic, antimony, bismuth, tellurium, polonium, astatine, lanthanum, cerium, pras odymium, neodymium, promethium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, thorium and oxidation states and isotopes thereof; in particular rubidium, magnesium, calcium, strontium, barium, yttrium, lutetium, chromium, manganese, technetium, rhenium, iron, ruthenium, osmium, cobalt, rhodium, nickel, palladium, platinum, copper, silver, gold, zinc, aluminium, gallium, indium, thallium, germanium, tin, antimony, bismuth samarium, europium, gadolinium, terbium, thorium and oxidation states or isotopes thereof; preferably magnesium (II), calcium (II), manganese (II), iron (II) and (III), ruthenium (II) and (III), nickel (II), palladium (II), platinum (II), copper (II), zinc (II), samarium (III), europium (III), terbium (III) or isotopes thereof.

For the present purpose, a particularly favourable test compound is a Cu^{2+}-phenanthroline complex, a Zn^{2+}-phenanthroline complex, a Mg^{2+}-phenanthroline complex, a Ca^{2+}-phenanthroline complex, a Cu^{2+}-bipyridyl complex, a Zn^{2+}-bipyridyl complex, a Mg^{2+}-bipyridyl complex, a Ca^{2+}-bipyridyl complex, a Cu^{2+}-1,4,8,11-tetraazacyclotetradecane, a Zn^{2+}-1,4,8,11-tetraazacyclotetradecane, a Mg^{2+}-1,4,8,11-tetraazacyclotetradecane, a Ca^{2+}-1,4,8,11-tetraazacyclotetradecane.

The method of the invention may suitably include a further step of determining, based on the three-dimensional structure of the specific protein in question (e.g. established by conventional X-ray crystallographic or NMR methods), or by the primary structure of the specific protein together with the three-dimensional structure of the class of proteins to which the specific protein belongs (e.g. established by sequence homology searches in DNA or amino acid sequence databases), the location of the metal ion binding amino acid residue in the mutated protein, and determining the location of at least one other amino acid residue in the vicinity of the metal ion binding amino acid residue. It is currently believed that in such a protein, the metal ion binding site may suitably be introduced by amino acid substitution 1) at a site where the protein binds to another protein, for example a regulatory protein; 2) at a site where the binding of the test compound will interfere with the cellular targeting of the protein; 3) at a site where the binding of the test compound will directly or indirectly interfere with the binding of substrate or the binding of an allosteric modulatory factor for the protein; 4) at a site where the binding of the test compound may interfere with the intra-molecular interaction of domains within the protein, for example the

interaction of a r gulatory domain with a catalytic domain; 5) at a site wher binding of the test compound will interfere with the folding of the protein, for example the folding of th protein into its active conformation; or 6) at a site which will control the activity of the protein, for example by an allosteric mechanism.

With respect to membrane proteins specifically, the method of the present invention may include a further step of determining, based on the primary structure of the specific membrane protein in question and the generic three-dimensional model of the class of membrane proteins to which the specific membrane protein belongs or is believed to belong, the location of the one metal ion binding amino acid residue in the mutated membrane protein, and determining the location of at least one other amino acid residue in the vicinity of the metal ion binding amino acid residue.

The term "in the vicinity of" is intended to include an amino acid residue located in the area defining the binding site of the metal ion chelate and at such a distance from the metal ion binding amino acid residue that it is possible, by attaching suitable functional groups to the test compound, to generate an interaction between said functional group or groups and said amino acid residue.

The term "generic three-dimensional model" is intended to indicate a model of the membrane protein established on the basis of, e.g., X-ray crystallographic data of a membrane protein of the same family, electron density maps of the membrane protein generated by cryo-electronmicroscopic analysis of two-dimensional membrane crystals (Baldwin, *EMBO J. 12*(4), 1993, pp. 1693-1703; Baldwin, *Curr. Opi. Cell. Biol. 6*, 1997, pp. 180-190; Herzyk et al. *J. Mol. Biol. 281*(4), 1998, pp. 741-754), knowledge of helical or other structures of the particular membrane protein, etc.

Such a determination may performed using a generic three-dimensional model of the membrane protein showing the spatial arrangement of the amino acid residues defining the area of the metal ion binding site, site-directed mutagenesis of a least one amino acid residue potentially involved in interaction with said metal ion chelated to the test compound, followed by expression of the mutated membrane protein in a suitable cell, contacting said cell or a portion thereof including the expressed mutated membrane protein with the test compound, and determining any effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal

transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent. If an amino acid residue involved in interaction with the metal ion is mutated to one which is not, this may be detected as a decrease in binding or other activity. In particular when the membrane protein is an orphan receptor, binding of a test compound is advantageously accomplished when the metal ion chelator portion of the test compound is detectable in itself, e.g. a fluorescent compound detectable by spectrophotometry, or labelled with a detectable labelling agent. Such labelling agents include, but are not limited to, radioactive isotopes or enzymes which may be detected by using suitable substrates, in accordance with conventional detection techniques.

Furthermore, the method of the invention may be used to identify at least one other amino acid residue in the vicinity of the metal chelating atom of the metal ion binding amino acid residue, said other amino acid residue being capable of binding at least one functional group of the test compound other than the metal ion.

Such a determination is performed using site-directed mutagenesis of a least one amino acid residue potentially involved in interaction with said functional group of the test compound other than the metal ion, followed by expression of the mutated membrane protein in a suitable cell, contacting said cell or a portion thereof including the mutated membrane protein with the test compound, and determining any effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent. If an amino acid residue involved in interaction with such a functional group of the test compound is mutated to one which is not, this may be detected as a decrease in binding or other activity.

Step (b) of the present method may advantageously comprise the further steps of improving the binding affinity of a heteroalkyl or heterocyclyl metal ion chelate to the mutated membrane protein, the method comprising

(i) selecting a heteroalkyl or heterocyclyl metal ion chelate with a binding affinity to the mutated membrane protein of 50 μM or better as identified by the screening method of claim 5,

(ii) mapping the site of the membrane protein to which the chelate binds using the method of claim 7,

(iii) locating at least one amino acid residue in the vicinity of the metal ion binding amino acid residue involved in interaction with at least one functional group of th chelate other than the metal ion,

(iv) providing the chelate with one or more functional groups of a nature and in a position optimised for interaction with said amino acid residue to generate a library of chelate derivatives,

(v) screening the derivatives of step (iv) by the method of claim 5, and

(vi) optionally repeating any one or a combination of two or more of steps (i)-(v) one or more times to generate heteroalkyl or heterocyclyl metal ion chelating compounds with an improved binding affinity for the mutated membrane protein.

The term "better" when used to indicate the concentration of test compound required to obtain a detectable binding to the membrane protein is intended to indicate a concentration below 50 μM. The concentration of the test compound required for binding to the membrane protein should preferably be lower than 10 μM, more particularly 1μM or lower, and more preferably in the nanomolar range such as 500 nM or lower.

Step (c) may comprise the further steps of generating a library of test compounds which are derivatives of a test compound found to interact with the wild-type membrane protein in step (c), and which lack a metal ion chelated thereto, each test compound in the library being provided with at least one functional group for interaction with at least one amino acid of the wild-type membrane protein, which functional group differs from at least one functional group of each of the other test compounds, and screening the test compound library for compounds interacting with the wild-type membrane protein. It is envisaged that it will be possible to improve the binding affinity of the test compounds used in step (c) of the present method in a similar fashion as that used to improve the binding affinity of compounds used in step (b), involving the following steps:

(i') selecting a test compound with a binding affinity to the membrane protein of 50 μM or better as identified by the screening method discussed above,

(ii') providing the test compound with one or more functional groups of a nature and in a position optimised for interaction with said amino acid residue to generate a library of test compound derivatives,

(iii') screening the derivatives of step (ii') by the method discussed above, and

(iv') optionally repeating any one or a combination two or more of steps (i')-(ii') one or more times to generate test compounds with an improved binding affinity for the membrane protein.

Step (i') may further comprise mapping the site of the membrane protein to which the test compound binds using the method discussed above, and determining at least one amino acid residue involved in interaction with at least one functional group of the test compound.

The screening in step (c) may be performed using effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent as discussed above.

Step (c) may also comprise the further step of determining, based on the primary structure of the specific membrane protein in question and the generic three-dimensional model of the class of membrane proteins to which the specific membrane protein belongs, and based on the information provided by the method of step (b) of the location of amino acid residues in the vicinity of the metal ion binding residue introduced in the mutated membrane protein, the location of an amino acid residue in the wild-type membrane protein binding at least one functional group of a test compound. The amino acid residue capable of binding at least one functional group of the test compound may advantageously be detected using site-directed mutagenesis of at least one amino acid residue of the wild-type membrane protein potentially involved in interaction with said functional group of the test compound, followed by expression of the mutated membrane protein in a suitable cell, contacting said cell or a portion thereof including the mutated membrane protein with the test compound, and determining any effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

The method of the invention is also expected to be useful for identifying compounds which interact with the membrane protein to either enhance or decrease its activity. Thus, the method may be used to identify agonists, partial agonists, positive modulators, antago-

nists, inv rse agonists or negative modulators of the membrane prot in in question, including orphan receptors.

A "partial agonist" is a compound which can only activate a membrane protein to a certain level compared to a full agonist. A "positive modulator" is a compound which is not able in itself to stimulate the activity of the membrane protein directly, but is able to do so indirectly through its impact on an agonist of the membrane protein. Such compounds are also known as enhancers.

An "inverse agonist" is a compound which is able to decrease the intrinsic, constitutive activity of a membrane protein. A "negative modulator" is a compound which is not in itself capable of inactivating the membrane protein, but is able to do so indirectly through its impact on an antagonist of inverse agonist of the membrane protein.

DESCRIPTION OF PREFERRED EMBODIMENTS

7TM receptors overview - A particularly interesting class of membrane proteins for use in the present invention is 7 transmembrane domain receptors (7TM receptors), also known as G-protein coupled receptors (GPCRs). This family of receptors constitutes the largest super-family of proteins in the human body and a large number of current drugs are directed towards 7TM receptors, for example: antihistamines (for allergy and gastric ulcer), beta-blockers (for cardiovascular diseases), opioids (for pain), and angiotensin antagonists (for hypertension). These current drugs are directed against relatively few receptors, which have been known for many years. To date, several hundred 7TMs have been cloned and characterized, and the total number of different types of 7TMs in humans is presumed to be between 1 and 2.000. The spectrum of ligands acting through 7TMs includes a wide variety of chemical messengers such as ions (e.g. calcium ions), amino acids (glutamate, γ-amino butyric acid), monoamines (serotonin, histamine, dopamine, adrenalin, noradrenalin, acetylcholine, cathecolamine, etc.), lipid messengers (prostaglandins, thromboxane, anandamide, etc.), purines (adenosine, ATP), neuropeptides (tachykinin, neuropeptide Y, enkephalins, cholecystokinin, vasoactive intestinal polypeptide, etc.), peptide hormones (angiotensin, bradykinin, glucagon, calcitonin, parathyroid hormone, etc.), chemokines (interleukin-8, RANTES, etc.), glycoprotein hormones (LH, FSH, TSH, choriogonadotropin, etc.) and proteases (thrombin). It is expected that a large number of the members of the 7TM superfamily of receptors will be suitable as drug tar-

gets. This notion is based on the fact that these receptors are involved in controlling major parts of the chemical transmission of signals between cells both in the endocrine and the paracrin system in the body as well as within the nervous system.

7TM receptor signalling - In 7TMs, binding of the chemical messenger to the receptor leads to the association of an intracellular G-protein which, in turn, is linked to a secondary messenger pathway. The G-protein consists of three subunits, an α subunit that binds and hydrolyses GTP, a β subunit and a γ subunit. When GDP is bound, the α subunit associates with the $\beta\gamma$ subunit to form an inactive heterotrimer that binds to the receptor. When the receptor is activated, a signal is transduced by a changed receptor conformation that activates the G-protein. This leads to the exchange of GDP for GTP on the α subunit, which subsequently dissociates from the receptor and the $\beta\gamma$ dimer, and activates downstream second messenger systems (e.g. adenylyl cyclase). The α subunit will activate the effector system until its intrinsic GTPase activity hydrolyses the bound GTP to GDP, thereby inactivating the α subunit. The $\beta\gamma$ subunit increases the affinity of the α subunit for GDP but may also be directly involved in intracellular signalling events.

Ligand-binding sites in general - Mutational analysis of 7TMs has demonstrated that functionally similar but chemically very different types of ligands can apparently bind in several different ways and still lead to the same function. Thus monoamine agonists appear to bind in a pocket relatively deep between TM-III, TM-V and TM-VI, while peptide agonists mainly appear to bind to the exterior parts of the receptors and the extracellular ends of the TMs (Strader et al., (1991) *J. Biol. Chem.* **266**, 5-8; Strader et al., (1994) *Ann. Rev. Biochem.* **63**, 101-132; Schwartz et al. *Curr. Pharm. Design.* (1995), **1**, 325-342). Moreover, ligands can be developed independent on the chemical nature of the endogenous ligand, for example non-peptide agonists or antagonists for peptide receptors. Such non-peptide antagonists for peptide receptors often bind at different sites from the peptide agonists of the receptors. For instance, non-peptide antagonists may bind in the pocket between TM-III, TM-V, TM-VI and TM-VII corresponding to the site where agonists and antagonists for monoamine receptors bind. It has been found that in the substance P receptor, when the binding site for a non-peptide antagonist has been exchanged for a metal ion binding site through introduction of His residues, no effect on agonist binding was observed (Elling et al., (1995) *Nature* **374**, 74-77; Elling et al., (1996) *EMBO J.* **15**, 6213-6219). It is believed that the non-peptide antagonist and the zinc ions act as antagonists by selecting and stabilizing an inactive conformation of the receptor that pre-

vents the binding and action of the agonist. This illustrates that drugs can b developed totally independent on knowledge of the endogenous ligand, since there need not b any overlap in their binding sites.

Generic numbering system for 7TMs - There are several families of 7TM receptors, which between each other are very divergent in respect of amino acid sequence. The largest family of 7TM receptors is composed of the rhodopsin-like receptors, which are named after the light-sensing molecule from our eye. Besides sharing an overall transmembrane domain composed of a seven helical bundle, many 7TMs also share other structural features such as the presence of a disulfide bridge between the top of TM-III and the middle of extracellular loop 3. Within the many hundred members of the rhodopsin-like receptor family, a number of residues are furthermore conserved - especially within each of the transmembrane segments; although only a single residue appear to be totally conserved in all signalling 7TM rhodopsin-like receptors, i.e. an Arg at the intracellular end of TM-III (ArgIII:26 – see below concerning the generic numbering nomenclature). Due to differences in length of especially the N-terminal segment, residues located at corresponding positions in different 7TM receptors are numbered differently in different receptors. However, based on the conserved key residues in each TM, a generic numbering system has been suggested (JM Baldwin, *EMBO J.* **12(4)**, 1993, pp. 1693-1703; TW Schwartz, *Curr. Opin. Biotech.* **5**, 1994, pp. 434-444), which is used here. In Fig. 1, a schematic depiction of the structure of rhodopsin-like 7TMs is shown with one or two conserved, key residues highlighted in each TM: AsnI:18; AspII:10; CysIII:01 and ArgIII:26; TrpIV:10; ProV:16; ProVI:15; ProVII:17. In relation to the present invention it is important that residues involved in for example metal ion binding sites can be described in this generic numbering system. For example, a tri-dentate metal ion site constructed in the tachykinin NK1 receptor (Elling et al., (1995) *Nature* **374**, 74-77) and subsequently transferred to the kappa-opioid receptor (Thirstrup et al., (1996) *J. Biol. Chem.* **271**, 7875-7878) and to the viral chemokine receptor ORF74 (Rosenkilde et al., *J. Biol. Chem.* 1999 Jan 8; 274(2), 956-61) can be described to be located between residues V:01, V:05, and VI:24 in all of these receptors although the specific numbering of the residues is very different in each of the receptors. It is only in the rhodopsin-like receptor family that a generic numbering system has been established; however, it should be noted that although the sequence identity between the different families of 7TM receptors is very low, it is believed that they may share a more-or-less common seven helical bundle structure. Thus,

all the techniques described in the present invention can be applied to the other families of 7TM rec ptors with minor modifications.

Metal ion sites in 7TMs - Naturally occurring metal ion sites have as yet only been described in two 7TM receptors, the tachykinin NK3 receptor (Rosenkilde et al., *FEBS Lett.* 1998 Nov 13; 439 (1-2), 35-40) and the galanin receptor (Kask et al., *EMBO J.* 1996 Jan 15; 15(2), 236). In the NK3 receptor Zn2+ was shown to act as an enhancer (positive modulator) for agonist binding and action without itself being an agonist. Through mutagenesis the metal ion binding site was mapped to residues V:01 and V:05 at the extracellular end of TM-V (Rosenkilde et al., *FEBS Lett.* 1998 Nov 13; 439(1-2), 35-40). In the galanin receptor zinc was shown to act as an antagonist for galanin binding, but the site was not characterized further (Kask et al., *EMBO J.* 1996 Jan 15; 15(2), 236). In contrast much work has been done on building up artificial metal ion sites in 7TM receptors (Elling et al., *Nature.* 1995 Mar 2; 374(6517), 74-7; Elling et al., *EMBO J.* 1996 Nov 15; 15(22), 6213-9; Elling et al., *Fold Des.* 1997; 2(4), S76-80; Elling et al., *Proc Natl Acad Sci USA* 1999 Oct 26; 96(22), 12322-7; Sheikh et al., *Nature.* 1996 Sep 26; 383(6598), 347-50). Based on this protein engineering work and on mutational analysis of ligand-binding sites as such at multiple locations in a number of wild-type 7TM receptors (Schwartz, T.W. (1994) *Curr.Opin.Biotech.* **5**, 434-444, Schwartz et al., *Curr. Pharm. Design*, 1995, **1**, 325-342) much experience has been gathered in relation to where metal ion sites can be located in the 7TM receptor structure as such. It is this knowledge which can be applied to wild-type, natural 7TM receptors to predict the occurrence of metal ion binding sites which have not previously been noted. Based on sequence analysis and molecular models it is found that such sites surprisingly occur in multiple 7TM receptors. Only few of these sites are probably addressed physiologically by free metal ions, for example when a receptor is expressed in brain regions where extracellular zinc concentrations may vary around 10^{-5} molar. Probably, many of the other previously unnoticed metal ion "sites" may just be a reflection of the fact that polar, metal ion binding amino acid residues (for example: His, Cys, Asp etc.) are frequently used by nature to face the water-exposed main ligand-binding crevice of 7TM receptors. These residues will in the current invention be used as initial attachment sites for metal ion chelating test compounds in the drug discovery process. It has recently been demonstrated that metal ion chelators could bind in artificially engineered metal ion binding sites with high affinity (Elling et al., *Proc Natl Acad Sci USA* 1999 Oct 26; 96(22), 12322-7). In relation to the present invention, the work on the engineered metal ion binding sites can be viewed as an indirect demonstration of the general

principl – however, performed by "reverse chemistry". That is, the metal ion binding site in the beta2-adrenergic receptor occurs in the mutated form of the receptor as opposed to the wild-type, and the mutations are used to "build up" the metal ion binding site and not to "map it". To illustrate this further, unpublished sequence analysis shows that the activating metal ion binding site between AspIII:08 and HisVII:06, which was built artificially into the beta2-adrenergic receptor, occurs naturally in the muscarinic M3 receptor and in the dopaminergic D4 receptor. Whether the III:08 to VII:06 site in these receptors in response to free metal ions or various metal ion chelates is activating or inactivating remains to be shown experimentally.

Use of metal ion chelate test compounds in mutated 7TM receptors in the initial phases of drug discovery - 7TM receptors are generally considered to be good targets in the established drug discovery and development process. That is, 7TMs often give a reasonable number of "hits" in for example high throughput screenings. However, some desired 7TM drug targets do not follow this rule and do not yield useful lead compound hits despite excessive search in various chemical screens. The present invention is aimed at eliminating this problem by creating artificial metal ion binding sites, which can serve as initial attachment site or initial site for molecular recognition for lead compounds being metal ion chelators where the metal ion can act as a guiding bridge between the chelator and the receptor. By mutating the receptor the gain is that a number of test compounds, metal ion chelators, will bind with useful affinity to the receptor, and the loss is that the compounds are binding to a mutated form and not the wild-type receptor. The test compound can then, based on detailed, precise knowledge about its binding site – a feature not normally known for lead compound or even high affinity drugs – be optimised in respect of general molecular recognition of the 7TM receptor target guided by molecular models of 7TM receptors. However, after systematically and rationally optimising the metal ion chelates to interact with a useful high affinity with other, opposing parts of the receptor, it will then be possible to use this newly established interaction as attachment site during a second round of chemical optimisation of the partly optimised lead compound to recognize the wild-type receptor. This second round of chemical optimisation can again be performed rationally guided by knowledge of the geometry and chemical nature of the residues, which initially had been mutated to create the metal ion binding site.

Orphan 7TM r ceptors - During th cloning of 7TM receptors many "extra" rec ptors were discovered for which no ligand was known, the so-called orphan receptors. Today there are several hundred of such orphan 7TM receptors. Based on characterization of their expression pattern in different tissues or expression during development or under particular physiological or patho-physiological conditions and based on the fact that the orphan receptors sequence-wise appear to belong to either established sub-families of 7TM receptors or together with other orphans in new families, it is believed that the majority of the orphan receptors are in fact important entities. As stated by representatives from the big pharmaceutical companies: Orphan 7TMs are "the next generation of drug targets" or "A neglected opportunity for pioneer drug discovery" (Wilson et al. Br.J.Pharmacol. 1998, 125: 1387-92; Stadel et al. Trends Pharmacol.Sci. 1997, 18: 430-37). Over the years ligands have been discovered for some of the orphan 7TM receptors, which then immediately have been recognized as "real" drug targets, for example: nocioceptin (for pain) (Reinscheid et al. Science 1995, 270: 792-94), orexin (for appetite regulation and regulation of energy homeostasis) (Sakurai et al. Cell 1998 92: 573-85), melanin-concentrating hormone (for appetite regulation) (Chambers et al. Nature 1999, 400: 261-65), and cysteinyl leukotrienes (inflammation, especially asthma) (Sarau et al. Mol. Pharmacol. 1999, 56: 657-63). In the latter case, a number of drugs (for example pranlukast, zafirlukast, montelukast, pobilukast) had in fact been developed in recent years against the receptor as a physiological entity without having access to the cloned receptor – which turned out to be a "well known" orphan receptor. The problem is that it is very difficult to characterize orphan receptors and find their endogenous ligands, since no assays are available for these receptors due to the lack of specific ligands – a real "catch 22" situation. The present invention is aimed at eliminating this problem. By introducing metal ion binding sites in orphan receptors at locations where it is known from previous work on multiple other 7TM receptors with known ligands and with binding and functional assays that binding of metal ions and metal ion chelates will act as either agonists or more common as antagonists, then it will be possible to establish binding assays and functional assays for the orphan receptors. Binding of metal ion chelates can be monitored either through functional assays in cases where agonistic metal ion sites are created, or through ligand binding assays. For example, many aromatic metal ion chelators are by themselves fluorescent and can therefore directly be used as tracers in binding assays. Or, radioactive or other measurable indicators can be incorporated into the metal ion chelator. By establishing a metal ion chelator based receptor analysis for the orphan receptors, it will become possible to search for the elusive endogenous ligands or it will be possible to use

the orphan receptors in various forms for drug discovery technology, for example high throughput screening.

It should be noted that due to the initial lack of knowledge of the endogenous ligand and therefore also lack of knowledge of the binding site for this ligand in the 7TM receptor, there is a certain danger that the introduced metal ion binding site can interfere with ligand binding or signal transduction. However, based on metal ion site engineering in multiple 7TM receptors and on mutational mapping of binding sites in multiple 7TM receptors, it will be possible to introduce such metal ion sites at different locations in the receptor in an attempt to eliminate this problem. Moreover, an artificial binding site and binding analysis, which may interfere with the binding of the natural ligand, may still be useful for screening for receptor ligands, for example antagonists.

CLAIMS

1. A method of identifying compounds acting as ligands of a membrane protein, the method comprising
(a) mutating a membrane protein or putative membrane protein by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the membrane protein;
(b) contacting the mutated membrane protein with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the membrane protein, and determining any binding of said test compound to said membrane protein; and
(c) contacting said membrane protein in wild-type non-mutated form, with at least one test compound determined to bind to the mutated membrane protein in step (b) but lacking a metal ion chelated thereto, and determining any interaction of the test compound lacking the metal ion with the wild-type membrane protein.

2. The method of claim 1, wherein, in step (a), the metal ion binding site is introduced in the ligand binding crevice of the membrane protein.

3. The method of claim 1, wherein, in step (a), the metal ion binding amino acid residue is introduced by site-directed mutagenesis.

4. A method according to claim 1, wherein step (b) comprises the further step of determining, based on the primary structure of the specific membrane protein in question and the generic three-dimensional model of the class of membrane proteins to which the specific membrane protein belongs, the location of the metal ion binding amino acid residue in the mutated membrane protein, and determining the location of at least one other amino acid residue in the vicinity of the metal ion binding amino acid residue.

5. A method according to claim 1, wherein the binding of the test compound to the mutated membrane protein in step (b) is determined using competition with binding of a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein, or using a metal ion chelator which is in itself detectable or labelled with a detectable labelling agent.

6. A method according to claim 1, wherein the amino acid residue in the vicinity of the metal ion binding amino acid residue is one which is capable of binding at least one functional group of the test compound other than the metal ion.

7. A method according to claim 6, wherein the amino acid residue capable of binding at least one functional group of the test compound other than the metal ion is detected using site-directed mutagenesis of at least one amino acid residue of the membrane protein potentially involved in interaction with said functional group of the test compound other than the metal ion, followed by expression of the mutated membrane protein in a suitable cell, contacting said cell or a portion thereof including the mutated membrane protein with the test compound, and determining any effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

8. A method of claim 1, wherein step (b) comprises the further steps of improving the binding affinity of a heteroalkyl or heterocyclyl metal ion chelate to the mutated membrane protein, the method comprising

(i) selecting a heteroalkyl or heterocyclyl metal ion chelate with a binding affinity to the mutated membrane protein of 50 μM or better as identified by the screening method of claim 5,

(ii) mapping the site of the membrane protein to which the chelate binds using the method of claim 7,

(iii) locating at least one amino acid residue in the vicinity of the metal ion binding amino acid residue involved in interaction with at least one functional group of the chelate other than the metal ion,

(iv) providing the chelate with one or more functional groups of a nature and in a position optimised for interaction with said amino acid residue to generate a library of chelate derivatives,

(v) screening the derivatives of step (iv) by the method of claim 5, and

(vi) optionally repeating any one or a combination of two or more of steps (i)-(v) one or more times to generate heteroalkyl or heterocyclyl metal ion chelating compounds with an improved binding affinity for the mutated membrane protein.

9. A method of claim 1, wherein step (c) comprises the further steps of generating a library of test compounds which are derivatives of a test compound found to interact with the wild-type membrane protein in step (c), each test compound in the library being provided with at least one functional group for interaction with at least one amino acid of the wild-type membrane protein, which functional group differs from at least one functional group of each of the other test compounds, and screening the test compound library for compounds interacting with the wild-type membrane protein.

10. A method of claim 1, wherein the screening in step (c) is performed using effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

11. A method according to claim 1, wherein step (c) comprises the further step of determining, based on the primary structure of the specific membrane protein in question and the generic three-dimensional model of the class of membrane proteins to which the specific membrane protein belongs, and based on the information provided by the method of claim 8 of the location of amino acid residues in the vicinity of the metal ion binding residue introduced in the mutated membrane protein, the location of an amino acid residue in the wild-type membrane protein binding at least one functional group of a test compound.

12. A method according to claim 11, wherein the amino acid residue capable of binding at least one functional group of the test compound is detected using site-directed mutagenesis of at least one amino acid residue of the wild-type membrane protein potentially involved in interaction with said functional group of the test compound, followed by expression of the mutated membrane protein in a suitable cell, contacting said cell or a portion thereof including the mutated membrane protein with the test compound, and determining any effect on binding in a competitive binding assay using a labelled ligand of the membrane protein, detection of any changes in signal transduction from the membrane protein or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

13. A method according to claim 9, comprising the further steps of
(i') selecting a test compound with a binding affinity to the membrane protein of 50 μM or better as identified by the screening method of claim 10,

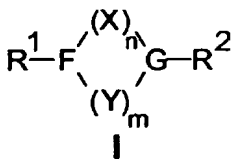
(ii') providing the test compound with one or mor functional groups of a natur and in a position optimised for interaction with said amino acid residue to generate a library of test compound derivatives,

(iii') screening the derivatives of step (ii') by the method of claim 10, and

(iv') optionally repeating any one or a combination two or more of steps (i')-(ii') one or more times to generate test compounds with an improved binding affinity for the membrane protein.

14. A method of claim 13, wherein step (i') further comprises mapping the site of the membrane protein to which the test compound binds using the method of claim 11, and determining at least one amino acid residue involved in interaction with at least one functional group of the test compound.

15. A method according to any of claims 1-14, wherein the test compound has the general formula I



I

wherein F is N, O or S and G is N, O or S;

and R^1 is an alkyl, aryl, cycloalkyl, alkoxy, ester, heteroalkyl, heterocycloalkyl or heteroaryl group, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide;

R^2 is an alkyl, aryl, cycloalkyl, alkoxy, ester, heteroalkyl or heteroaryl group, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide;

R^1 and/or R^2 optionally forming a fused ring together with the heteroatom to which it is attached, or R^1 and/or R^2 optionally forming a fused ring together with the heteroatom to which it is attached and a carbon atom of X or Y;

X is $-CH_2-$, CH_2-CH_2-, $-CH_2-S-CH_2-$, $-CH_2-N-CH_2-$, $-CH=CH-CH=CH-$, $-CH_2-(Z)_a-(P)_c-(W)_b-CH_2-$, $-CH_2-O-CH_2-$, wherein

each of Z and W are independently C, S, O or N, and

P is -CH- or $-CH_2-$;

Y is $-CH_2-$, CH_2-CH_2-, $-CH_2-S-CH_2-$, $-CH_2-N-CH_2-$, $-CH=CH-CH=CH-$, $-CH_2-(Z)_a-(P)_c-(W)_b-CH_2-$, $-CH_2-O-CH_2-$, wher in

each of Z and W are ind p ndently C, S, O or N, and

P is -CH- or $-CH_2-$;

n is 0 or an integer of 1-5,

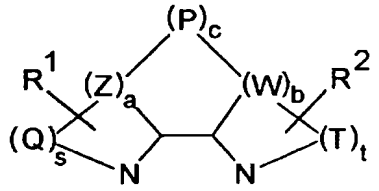
m is 0 or an integer of 1-5,

a is an integer of 1-3,

b is an integer of 1-3, and

c is an integer of 1-7.

16. A method according to claim 15, wherein the test compound has the general formula III

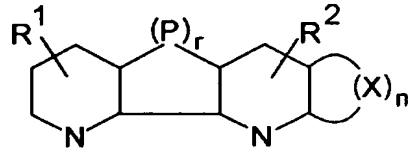


III

wherein R^1, R^2, Z, W, P, a, b and c are as indicated above, and

each of Q and T is independently -CH- or $-CH_2-$, s is an integer of 1-7, and t is an integer of 1-7.

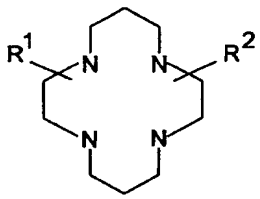
17. A method according to claim 15, wherein the test compound has the general formula IV



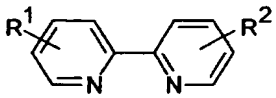
IV

wherein R^1, R^2, P, X and n are as indicated above, and r is an integer of 1-3.

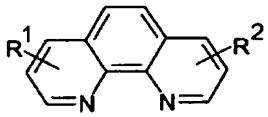
18. A method according to claim 15, wherein the test compound is select d from a compound of formula IIa, IIb or IIc



IIa



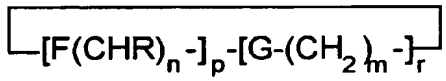
IIb



IIc

wherein R^1 and R^2 are as indicated above in formula I.

19. A method according to claim 15, wherein the test compound has the general formula V



V

wherein F is N,O or S and G is N,O or S,
n is an integer from 1 to 5,
m is 0 or an integer from 1 to 5,
p is 0 or an integer from 1 to 8,
r is 0 or an integer from 1 to 8, and
R is alkyl, aryl, cycloalkyl, alkoxy, heteroalkyl or heteroaryl, optionally substituted with one or more substituents selected from halogen, nitro, cyano, amino, alkyl, alkoxy, carboxy, amide or sulfonamide.

20. A method according to claim 1, wherein the membrane protein is an integral membrane protein.

21. A method according to claim 20, wherein the membrane protein comprises 1-14 transmembrane domains.

22. A method according to claim 21, wherein the membrane protein is a receptor such as a tyrosine kinase receptor, e.g. a growth factor receptor such as the growth hormone, insulin, epidermal growth factor, transforming growth factor, rythropoietin, colony-stimulating factor, platelet-derived growth factor receptor or nerve growth factor receptor (TrkA or TrkB).

23. A method according to claim 21, wherein the membrane protein is a purinergic ion channel.

24. A method according to claim 21, wherein the membrane protein is a ligand-gated ion channel, such as a nicotinic acetylcholine receptor, GABA receptor, or glutamate receptor (NMDA or AMPA).

25. A method according to claim 21, wherein the membrane protein is a voltage-gated ion channel, such as a potassium, sodium, chloride or calcium channel.

26. A method according to claim 21, wherein the membrane protein is a G-protein coupled receptor, such as the acetylcholine, adenosine, adrenoceptors, anaphylatoxin chemotactic receptor, angiotensin, bombesin (neuromedin), bradykinin, calcitonin, calcitonin gene related peptide, conopressin, corticotropin releasing factor, amylin, adrenomedullin, calcium, cannabinoid, chemokine, cholecystokinin, conopressin, corticotropin-releasing factor, dopamine, eicosanoid, endothelin, fMLP, $GABA_B$, galanin, gastrin, gastric inhibitory peptide, glucagon, glucagon-like, glutamate, glycoprotein hormone (e.g. FSH, LSH, TSH, LH), gonadotropin releasing hormone, growth hormone releasing hormone, growth hormone secretagogue, histamine, 5-hydroxytryptamine, leukotriene, lysophospholipid, melanocortin, melatonin, neuropeptide Y, neurotensin, olfactory, opioid, opioid-like, opsin, orexin, oxytocin, parathyroid hormone/parathyroid hormone-related, P2Y, pheromone, platelet-activating factor, prostanoid, protease-activated, secretin, somatostatin, tachykinin, thrombin and protease activated, thyrotropin-releasing hormone, pituitary adenylate activating peptide, vasopressin, vasoactive intestinal peptide and viral encoded receptors; in particular galanin, P2Y, chemokine, melanocortin, bombesin, cannabinoid, lysophospholipid, fMLP, neuropeptide Y, tachykinin, dopamine, histamine, 5-hydroxytryptamine, histamine, mas-proto-oncogene, acetylcholine, oxytocin, herpes virus encoded, epstein-barr virus encoded, cytomegalovirus encoded and bradykinin receptors; preferably galanin receptor type 1, leukotriene B4 receptor, CXCR5, melanocortin-1 receptor, melanocortin-3 recep-

tor, melanocortin-4 receptor, melanocortin-5 receptor, bombesin receptor subtype 3, cannabinoid rec ptor 1, cannabinoid receptor 2, EDG-2, EDG-4, FMLP-related receptor I, FMLP-related receptor II, Y6 receptor, Y5 receptor, NK-1 receptor, NK-3 receptor, D2 receptor (short), D2 receptor (long), D4 receptor (D2C), duffy antigen, Q89609 herpes virus type 2, VU51_HSV6U, histamine H1 receptor, MAS proto-oncogene, muscarinic M1 receptor, muscarinic M2 receptor, muscarinic M3 receptor, muscarinic M5 receptor, oxytocin receptor, XCR1 receptor, EBI2 receptor, RDC1 receptor, O76000 receptor, O15218 receptor, GPR12 receptor or GPR3 receptor.

27. A method according to claim 21, wherein the membrane protein is a transporter protein, such as a GABA, monoamine or nucleoside transporter.

28. A method according to claim 21, wherein the membrane protein is a multidrug resistance protein, e.g. a P-glycoprotein, multidrug resistance associated protein, drug resistance associated protein, lung resistance related protein, breast cancer resistance protein, adenosine triphosphate-binding cassette protein, Bmr, QacA or EmrAB/TolC pump.

29. A method according to claim 21, wherein the membrane protein is a cell adhesion molecule, e.g. NCAM, VCAM or ICAM.

30. A method according to claim 21, wherein the membrane protein is an enzyme such as adenylyl cyclase.

31. A method according to claim 20, wherein the membrane protein is an orphan receptor.

32. A method according to any of claims 1-31, wherein the test compound is an agonist, partial agonist or positive modulator of the membrane protein.

33. A method according to any of claims 1-31, wherein the test compound is an antagonist, inverse agonist or negative modulator of the membrane protein.

34. A method of preferentially stabilising a membrane protein in an active conformation, the method comprising contacting said membrane protein with an effective amount of a compound as defined in any of claims 15-19 acting as an agonist, partial agonist or positive modulator of the membrane protein.

35. A method of preferentially stabilising a membrane protein in an inactive conformation, the method comprising contacting said membrane protein with an effective amount of a compound according to any of claims 15-19 acting as an antagonist, inverse agonist or negative modulator of the membrane protein.

36. A method of identifying a compound acting as a ligand of an orphan receptor, the method comprising

(a) mutating an orphan receptor or putative orphan receptor by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the orphan receptor;
(b) contacting the mutated orphan receptor with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the orphan receptor, and determining any binding of said test compound to said orphan receptor.

37. A method according to claim 36, wherein the binding of the test compound to the metal ion binding site of the orphan receptor is determined by using a metal ion chelator which is itself detectable or labelled with a detectable labelling agent.

38. A method according to claim 36, further comprising contacting the mutated orphan receptor with one or more test substances which are not metal ion chelates as well as contacting the mutated orphan receptor with a test compound found to bind to the mutated orphan receptor in step (b), and determining any binding of such test substance or test substances as altered binding of the metal ion chelate to the orphan receptor.

39. A method according to claim 36, wherein step (b) comprises the further step of determining, based on the primary structure of the specific orphan receptor in question and the generic three-dimensional model of the class of receptors to which the specific orphan receptor belongs or is predicted to belong, the location of the metal ion binding amino acid site in the mutated orphan receptor, and determining the location of at least one other amino acid residue in the vicinity of the metal ion binding amino acid residue.

40. A method according to any of claims 36-39, wherein the metal chelator moi ty of the test compound is as indicated in any of claims 15-19.

41. A method of screening for compounds capable of acting as ligands on an orphan receptor, the method comprising contacting an orphan receptor or putative orphan receptor which is mutated by substituting at least one amino acid residue by at least one other amino acid residue capable of binding metal ions, resulting in the introduction of a metal ion binding site into the orphan receptor, with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue capable of binding metal ions, as well as one or more test substances which are not metal ion chelates under conditions permitting binding of the test compound and one or more test substances to the mutated orphan receptor, and determining any binding of said test substance or substances to said mutated orphan receptor by measuring altered binding of the metal chelate test compound to the receptor or changes in signal transduction from the orphan receptor.

42. A method of identifying a compound capable of interacting with a protein so as to modify its biological activity, the method comprising

(a) mutating a protein by substituting at least one amino acid residue by at least one other amino acid residue capable of binding a metal ion, resulting in the introduction of a metal ion binding site into the protein;
(b) contacting the mutated protein with a test compound which comprises a moiety including at least one heteroatom for chelating a metal ion and further comprising a metal ion chelated thereto for binding to the amino acid residue introduced in step (a) under conditions permitting binding of the test compound to the protein, and determining any effect of the binding of said test compound to said protein by detecting any change in the biological activity of said protein; and
(c) contacting said protein in wild-type, non-mutated form with at least one test compound determined to bind the mutated protein in step (b) but lacking a metal ion chelated thereto, and determining any interaction of the test compound lacking the metal ion with the wild-type protein.

43. The method of claim 42, wherein, in step (a), the metal ion binding site is introduced by amino acid substitution 1) at a site where the protein binds to another protein, for ex-

ample a r gulatory protein; 2) at a site wher the binding of the test compound will interfere with the cellular targeting of the protein; 3) at a site where the binding of the test compound will directly or indirectly interfere with the binding of substrate or the binding of an allosteric modulatory factor for the protein; 4) at a site where the binding of the test compound may interfere with the intra-molecular interaction of domains within the protein, for example the interaction of a regulatory domain with a catalytic domain; 5) at a site where binding of the test compound will interfere with the folding of the protein, for example the folding of the protein into its active conformation; or 6) at a site which will control the activity of the protein, for example by an allosteric mechanism.

44. The method of claim 42, wherein, in step (a), the metal ion binding amino acid residue is introduced by site-directed mutagenesis.

45. A method according to claim 42, wherein step (b) comprises the further step of determining, based on the three-dimensional structure of the specific protein in question or the primary structure of the specific protein together with a three-dimensional model of the class of proteins to which the specific protein belongs, the location of the metal ion binding amino acid residue in the mutated protein, and determining the location of at least one other amino acid residue in the vicinity of the metal ion binding amino acid residue.

46. A method according to claim 42, wherein the binding of the test compound to the mutated protein in step (b) is determined using detection of any changes in the biological activity of the protein, competition with binding of a labelled ligand of the protein, or using a metal ion chelator which is in itself detectable or labelled with a detectable labelling agent.

47. A method according to claim 45, wherein the amino acid residue in the vicinity of the metal ion binding amino acid residue is one which is capable of binding at least one functional group of the test compound other than the metal ion.

48. A method according to claim 47, wherein the amino acid residue capable of binding at least one functional group of the test compound other than the metal ion is detected using site-directed mutagenesis of at least one amino acid residue of the protein potentially involved in interaction with said functional group of the test compound other than the metal ion, followed by expression of the mutated protein in a suitable cell, contacting said cell or a portion thereof including the mutated protein with the test compound, and detecting any

changes in the biological activity of the protein, determining any effect on binding in a competitive binding assay using a labelled ligand of the protein, or using a chelating agent which is in itself detectabl or labelled with a detectable labelling agent.

49. The method of claim 42, wherein step (b) comprises the further steps of improving the binding affinity of a heteroalkyl or heterocyclyl metal ion chelate to the mutated protein, the method comprising

(i) selecting a heteroalkyl or heterocyclyl metal ion chelate with a binding affinity to the mutated protein of 50 μM or better as identified by the screening method of claim 46,

(ii) mapping the site of the protein to which the chelate binds using the method of claim 45 and/or 48,

(iii) locating at least one amino acid residue in the vicinity of the metal ion binding amino acid residue involved in interaction with at least one functional group of the chelate other than the metal ion,

(iv) providing the chelate with one or more functional groups of a nature and in a position optimised for interaction with said amino acid residue to generate a library of chelate derivatives,

(v) screening the derivatives of step (iv) by the method of claim 46, and

(vi) optionally repeating any one or a combination of two or more of steps (i)-(v) one or more times to generate heteroalkyl or heterocyclyl metal ion chelating compounds with an improved binding affinity for the mutated protein.

50. A method of claim 42, wherein step (c) comprises the further steps of generating a library of test compounds which are derivatives of a test compound found to interact with the wild-type protein in step (c), each test compound in the library being provided with at least one functional group for interaction with at least one amino acid of the wild-type protein, which functional group differs from at least one functional group of each of the other test compounds, and screening the test compound library for compounds interacting with the wild-type protein.

51. A method of claim 42, wherein the screening in step (c) is performed by detecting any changes in the biological activity of the protein, detecting an effect on binding in a competitive binding assay using a labelled ligand of the protein, or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

52. A method according to claim 42, wherein step (c) comprises the further step of determining, bas d on th thr -dimensional structure of th specific protein in question or the primary structure of the specific protein together with the three-dimensional model of the class of proteins to which the specific protein belongs, and based on the information provided by the method of claim 49 of the location of amino acid residues in the vicinity of the metal ion binding residue introduced in the mutated protein, the location of an amino acid residue in the wild-type protein binding at least one functional group of a test compound.

53. A method according to claim 52, wherein the amino acid residue capable of binding at least one functional group of the test compound is detected using site-directed mutagenesis of at least one amino acid residue of the wild-type protein potentially involved in interaction with said functional group of the test compound, followed by expression of the mutated protein in a suitable cell, contacting said cell or a portion thereof including the mutated protein with the test compound, and determining any effect on binding using detection of any changes in the biological activity of the protein, a competitive binding assay using a labelled ligand of the protein, or using a chelating agent which is in itself detectable or labelled with a detectable labelling agent.

54. A method according to claim 42, wherein step (c) comprises the further steps of
(i') selecting a test compound with a binding affinity to the protein of 50 μM or better as identified by the screening method of claim 51,
(ii') providing the test compound with one or more functional groups of a nature and in a position optimised for interaction with said amino acid residue to generate a library of test compound derivatives,
(iii') screening the derivatives of step (ii') by the method of claim 51, and
(iv') optionally repeating any one or a combination two or more of steps (i')-(ii') one or more times to generate test compounds with an improved binding affinity for the protein.

55. A method of claim 54, wherein step (i') further comprises mapping the site of the protein to which the test compound binds using the method of claim 52, and determining at least one amino acid residue involved in interaction with at least one functional group of the test compound.

56. The method of any of claims 42-55, wherein the metal chelator moiety of the test compound is as indicated in any of claims 15-19.

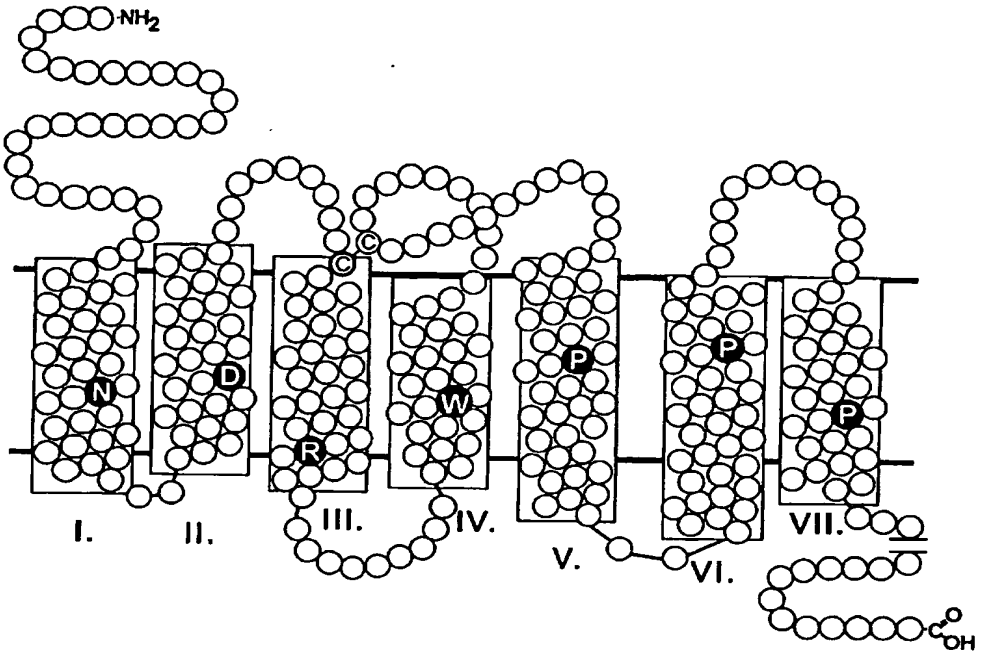


Fig. 1